



Universidad Autónoma del Estado de México

Material didáctico multimedia

Sólo visión

Título del material: La metodología CRISP-DM

Unidad de Aprendizaje Minería de Datos
Licenciatura de Ingeniería en Computación
Facultad de Ingeniería

Elaborado por M en I Sara Vera Noguez
Durante el período intensivo verano 2016

Contenido

No se encontraron entradas de tabla de contenido.

Justificación académica

Introducción

El curso de Minería de Datos corresponde a una Unidad de Aprendizaje (L41093) de la línea de acentuación de desarrollo de software, de la licenciatura de Ingeniería en Computación, que se imparte en la Facultad de Ingeniería, así como en los Centros Universitarios: Atlacomulco, Ecatepec, Texcoco, Valle de Chalco, Valle de México, Valle de Teotihuacán y Zumpango.

Se trata de un curso teórico - práctico de 6 créditos, con una carga de 2 horas teóricas y horas prácticas a la semana.

Competencias

De acuerdo a lo plasmado en el programa de estudios vigente y avalado por los HH Consejos, las competencias a desarrollar en el curso son las siguientes:

El alumno:

- Conocerá las bases metodológicas para la extracción de conocimiento a partir de grandes repositorios de información.

- Conocerá las bases teóricas de las técnicas de minería más usuales que permiten la obtención de modelos y patrones a partir de la vista minable.
- Conocerá el software disponible para realizar minería de datos.
- Podrá conjuntar los conocimientos adquiridos para realizar un proyecto de minería de datos en un contexto real.
- Podrá interpretar y evaluar los resultados obtenidos para trasladarlos en acciones recomendadas.
- Conocerá la metodología CRISP-DM y la minería de datos complejos.

Estructura del curso

De acuerdo al programa actual oficial, la unidad de aprendizaje se encuentra estructurada por las siguientes cinco unidades de aprendizaje:

1. Introducción a la minería de datos
2. La minería de datos en el proceso de KDD (Knowledge Discovery in Databases)
3. Tareas y métodos en minería de datos
4. Herramientas de software disponibles para minería de datos
5. Metodología CRISP-DM y minería de datos complejos

Importancia del material en el contexto del curso

Ya que la Unidad es parte de la línea de acentuación de desarrollo de software, resulta de gran importancia que los estudiantes identifiquen cómo las tareas de minería de datos se mezclan en el proceso de desarrollo de software, y esto se puede lograr revisando la metodología CRISP-DM, a la par que se recuerdan dos de las metodologías ampliamente utilizadas en desarrollo de software, con las que los estudiantes de esta línea de acentuación se encuentran ampliamente familiarizados: Up y XP.

La intención de retomar a metodologías conocidas, para a partir de ellas entender y relacionar las tareas de la nueva metodología; obedece a las ideas de la zona de desarrollo próximo Vigotsky, mencionada por varios autores como Moll.

Así pues se busca acceder al nuevo conocimiento, el relativo a la metodología CRISP-DM, a partir de conocimiento previo, permitiendo vincular el contenido de este curso con el de otros.

Recomendaciones de uso

Si bien se trata de contenidos relacionados con la última unidad, lo relativo a la metodología CRISP-DM se aconseja abordarlo al terminar la unidad 1, antes de comenzar con el contenido de la unidad 2, el proceso KDD, ya que estrictamente este proceso debería estar inmerso en el marco de la metodología, y de esta forma se ayuda a que los estudiantes partan de algo conocido y general, como lo es el proceso de desarrollo de software, para ir hacia el contenido nuevo y específico de la Minería de datos.

El material contiene información no sólo tomada de fuentes del área de minería de Datos, sino también de la de Ingeniería de software, conjuntando información útil en punto.

Se aconseja proyectar el material a la vez que se da la clase y plantean algunas actividades que lleven al estudiante a reflexionar en cuanto a la semejanza de la metodología CRISP-DM con UP y XP.

Este contenido puede ser abordado en una sesión de clase.

Bibliografía

Biggs, J., & Biggs, J. B. (2005). "Calidad del aprendizaje universitario" (Vol. 7). España: Narcea ediciones.

De la Rosa Reyes, María de los Ángeles (2004), El Desarrollo de Competencias Comunicativas: uno de los Principales Retos en la Educación Superior". México, UNAM.

Demarchi, Claudia. (2009). "Estrategias de comunicación en entornos virtuales". Argentina: Escuela de Ciencias de la Información

Han, D. J. (2007). Principles of Data Mining, EEUU: MIT Press.

Hernández Orallo, J., M. J. Ramírez Quintana, et al. (2004). Introducción a la Minería de Datos. España, Pearson Educación SA.

Introducción a la Minería de Datos. España, Pearson Educación SA.

Maimon, O. Z. and L. Rokach (2005). Data mining and knowledge discovery handbook. USA, Springer.

Marroquín Pérez, Manue. (1995). "Importancia de la Comunicación Interpersonal en *La Comunicación Interpersonal*". España: Ediciones Mensajero

Moll, Luis (). "La zona de desarrollo próximo de Vigotsky: Una consideración de sus implicaciones para la enseñanza". EEUU, Universidad de Arizona

Pérez López, C. and D. Santín Gonzalez (2006). Data Mining- Soluciones Con Enterprise Miner. México, Alfaomega, Ra-Ma.

Sumathi, S. and S. N. Sivanandam (2006). Introduction to data mining and its applications. Berlín, Germany, Springer-Verlag New York Inc.

Tan, P. N., M. Steinbach, et al. (2005). Introduction to data mining, Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA. Maimon, O. Z. and L. Rokach (2005). Data mining and knowledge discovery handbook. USA, Springer.

Whitten, Bentley, (2008). "Análisis de sistemas, diseño y métodos" ,7 a edición, McGraw Hill,.